Testing spatial structures 000000000 000000 00000 Multivariate analysis and spatial patterns

Inferring directionality 00000

Spatial genetics: a methodological overview

T. Jombart

Imperial College London

Bogota 02-12-2010

Testing spatial structures 000000000 000000 00000 Multivariate analysis and spatial patterns

Inferring directionality 00000

Outline

Introduction

Testing spatial structures

Moran's Index Mantel's correlation Correlogram

Multivariate analysis and spatial patterns

Inferring directionality

Testing spatial structures 000000000 000000 00000 Multivariate analysis and spatial patterns

Inferring directionality 00000

Outline

Introduction

Testing spatial structures Moran's Index

- Mantel's correlation
- Correlogram

Multivariate analysis and spatial patterns

Inferring directionality

Testing spatial structures 00000000 000000 00000 Multivariate analysis and spatial patterns

Inferring directionality 00000

Island model

Reproduction within populations + migration.





Multivariate analysis and spatial patterns 000000000

Inferring directionality 00000

Hierarchical island model

Reproduction within subpopulations + stratified migration.





Multivariate analysis and spatial patterns 000000000

Inferring directionality 00000

Isolation by distance (IBD)

Reproduction between neighbours \rightarrow 'diffusion' of genes





Multivariate analysis and spatial patterns

Inferring directionality 00000

Inbreeding avoidance

Mating with individuals from another population \rightarrow 'repulsion' structure



Multivariate analysis and spatial patterns

Inferring directionality 00000

Genetic models and spatial structures

- *island / hierarchical island model*: patches of related genotypes
- isolation by distance (IBD): clines of genetic differentiation
- inbreeding avoidance: repulsion structure

Multivariate analysis and spatial patterns

Inferring directionality 00000

Genetic models and spatial structures

- *island / hierarchical island model*: patches of related genotypes
- isolation by distance (IBD): clines of genetic differentiation
- *inbreeding avoidance*: repulsion structure

Multivariate analysis and spatial patterns

Inferring directionality 00000

Genetic models and spatial structures

- *island / hierarchical island model*: patches of related genotypes
- isolation by distance (IBD): clines of genetic differentiation
- *inbreeding avoidance*: repulsion structure

Multivariate analysis and spatial patterns

Inferring directionality 00000

Genetic models and spatial structures

- *island / hierarchical island model*: patches of related genotypes
- isolation by distance (IBD): clines of genetic differentiation
- *inbreeding avoidance*: repulsion structure

Multivariate analysis and spatial patterns

Inferring directionality 00000

- test the existence of spatial patterns
- detect the scales of spatial structures
- describe spatial genetic structures
- infer directionality of the patterns (migrations)

Multivariate analysis and spatial patterns

Inferring directionality 00000

- test the existence of spatial patterns
- detect the scales of spatial structures
- describe spatial genetic structures
- infer directionality of the patterns (migrations)

Multivariate analysis and spatial patterns

Inferring directionality 00000

- test the existence of spatial patterns
- detect the scales of spatial structures
- describe spatial genetic structures
- infer directionality of the patterns (migrations)

Multivariate analysis and spatial patterns

Inferring directionality 00000

- test the existence of spatial patterns
- detect the scales of spatial structures
- describe spatial genetic structures
- infer directionality of the patterns (migrations)

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Outline

Introduction

Testing spatial structures Moran's Index

Mantel's correlation

Correlogram

Multivariate analysis and spatial patterns

Inferring directionality

Multivariate analysis and spatial patterns

Inferring directionality 00000

Outline

Introduction

Testing spatial structures Moran's Index

Mantel's correlation Correlogram

Multivariate analysis and spatial patterns

Inferring directionality

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Spatial autocorrelation

Definitions:

- *in general*: values of a variable non independent from the corresponding spatial locations
- *in genetics*: genetic distance is correlated to spatial distance

Two types of spatial autocorrelation:

- *positive*: closer individuals are more similar than at random
- negative: closer individuals are more dissimilar than at random

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Spatial autocorrelation

Definitions:

- *in general*: values of a variable non independent from the corresponding spatial locations
- *in genetics*: genetic distance is correlated to spatial distance

Two types of spatial autocorrelation:

- *positive*: closer individuals are more similar than at random
- negative: closer individuals are more dissimilar than at random

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Spatial autocorrelation: illustration



Testing spatial structures 00000000 000000 00000 Multivariate analysis and spatial patterns

Inferring directionality 00000

Spatial weights and the lag vector

Matrix of spatial weights ${\bf L}$

Row i : uniform weights for neighbours of i.



	a	b	C	d	e
а	0.000	0.500	0.000	0.000	0.500
b	0.333	0.000	0.000	0.333	0.333
С	0.000	0.000	0.000	1.000	0.000
d	0.000	0.333	0.333	0.000	0.333
е	0.333	0.333	0.000	0.333	0.000

Let \mathbf{x} be a variable with one value at each location.

The lag vector Lx computes mean values of neighbours.

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Spatial weights and the lag vector

Matrix of spatial weights ${\bf L}$

Row i: uniform weights for neighbours of i.



а	b	с	d	e
0.000	0.500	0.000	0.000	0.500
0.333	0.000	0.000	0.333	0.333
0.000	0.000	0.000	1.000	0.000
0.000	0.333	0.333	0.000	0.333
0.333	0.333	0.000	0.333	0.000
	a 0.000 0.333 0.000 0.000 0.333	a b 0.000 0.500 0.333 0.000 0.000 0.000 0.000 0.333 0.333 0.333	a b c 0.000 0.500 0.000 0.333 0.000 0.000 0.000 0.000 0.000 0.000 0.333 0.333 0.333 0.333 0.000	a b c d 0.000 0.500 0.000 0.000 0.333 0.000 0.000 0.333 0.000 0.000 0.000 1.000 0.333 0.333 0.300 0.333 0.000 0.333 0.333 0.333

Let \mathbf{x} be a variable with one value at each location.

The lag vector Lx computes mean values of neighbours.

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Spatial weights and the lag vector

Matrix of spatial weights ${\bf L}$

Row i: uniform weights for neighbours of i.



а	Ь	С	d	e
0.000	0.500	0.000	0.000	0.500
0.333	0.000	0.000	0.333	0.333
0.000	0.000	0.000	1.000	0.000
0.000	0.333	0.333	0.000	0.333
0.333	0.333	0.000	0.333	0.000
	a 0.000 0.333 0.000 0.000 0.333	a b 0.000 0.500 0.333 0.000 0.000 0.000 0.000 0.333 0.333 0.333	a b c 0.000 0.500 0.000 0.333 0.000 0.000 0.000 0.000 0.000 0.000 0.333 0.333 0.333 0.333 0.000	a b c d 0.000 0.500 0.000 0.000 0.333 0.000 0.000 0.333 0.000 0.000 0.000 1.000 0.333 0.333 0.300 0.333 0.000 0.333 0.333 0.333

Let \mathbf{x} be a variable with one value at each location.

The lag vector Lx computes mean values of neighbours.

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

A variable and its lag-vector

Lag vector :

Random:



Random:

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

A variable and its lag-vector

Lag vector :





	Df	Sum Sq	Mean Sq	F value	Pr(>F)
x	1	0.02	0.02	0.06	0.8081
Residuals	98	31.53	0.32		



Positive

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

A variable and its lag-vector

Lag vector :





	Df	Sum Sq	Mean Sq	F value	Pr(>F)
×G	1	65.91	65.91	245.69	0.0000
Residuals	98	26.29	0.27		

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

A variable and its lag-vector

Lag vector :







-	Df	Sum Sq	Mean Sq	F value	Pr(>F)
×L	1	87.56	87.56	77.80	0.0000
Residuals	98	110.29	1.13		

Multivariate analysis and spatial patterns

Inferring directionality 00000

Moran's index: definition

Moran's *I*:

$$I(\mathbf{x}) = \frac{\mathbf{x}^T \mathbf{L} \mathbf{x}}{n} \frac{1}{\mathsf{var}(\mathbf{x})}$$

where:

- $\mathbf{x} \in \mathbb{R}^n$: a centred variable (e.g. allele frequency, PC)
- L : matrix of spatial weights (nxn)
- Lx : lag vector
- $I_0 = \frac{-1}{n-1} \approx 0$: null value (no autocorrelation, i.e. random spatial distribution)
- \Rightarrow Moran's I varies like $\langle \mathbf{x}, \mathbf{Lx} \rangle$.

Multivariate analysis and spatial patterns

Inferring directionality 00000

Moran's index: definition

Moran's *I*:

$$I(\mathbf{x}) = \frac{\mathbf{x}^T \mathbf{L} \mathbf{x}}{n} \frac{1}{\mathsf{var}(\mathbf{x})}$$

where:

- $\mathbf{x} \in \mathbb{R}^n$: a centred variable (e.g. allele frequency, PC)
- L : matrix of spatial weights (nxn)
- Lx : lag vector
- $I_0 = \frac{-1}{n-1} \approx 0$: null value (no autocorrelation, i.e. random spatial distribution)
- \Rightarrow Moran's I varies like $\langle \mathbf{x}, \mathbf{Lx} \rangle$.

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Variable, lag-vector, Moran's I

Lag vector :

Random:



Moran's *I*:

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Variable, lag-vector, Moran's I

Lag vector :



Moran's I: $I(\mathbf{x}) \approx I_0$

Random:



17/49

Positive autocorrelation:

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Variable, lag-vector, Moran's I

Lag vector :



Moran's I: $I(\mathbf{x}) > I_0$



Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Variable, lag-vector, Moran's I

Lag vector :

Negative autocorrelation:



Moran's I: $I(\mathbf{x}) < I_0$

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Testing Moran's I

Monte Carlo procedure:

- compute *I* from the data
- permute randomly the locations to get a value of *I* under *H*₀: "x is distributed at random across space."
- repeat this operation a large number of times to obtain a reference distribution of I under ${\cal H}_0$
- compare initial value to the reference distribution to get a p-value.

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Testing Moran's I

Monte Carlo procedure:

- compute *I* from the data
- permute randomly the locations to get a value of *I* under *H*₀: "**x** is distributed at random across space."
- repeat this operation a large number of times to obtain a reference distribution of I under ${\cal H}_0$
- compare initial value to the reference distribution to get a p-value.

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Testing Moran's I

Monte Carlo procedure:

- compute *I* from the data
- permute randomly the locations to get a value of *I* under *H*₀: "x is distributed at random across space."
- repeat this operation a large number of times to obtain a reference distribution of I under ${\cal H}_0$
- compare initial value to the reference distribution to get a p-value.
Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Testing Moran's I

- compute *I* from the data
- permute randomly the locations to get a value of *I* under *H*₀: "x is distributed at random across space."
- repeat this operation a large number of times to obtain a reference distribution of I under ${\cal H}_0$
- compare initial value to the reference distribution to get a p-value.

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality

Application: testing spatial structures in principal components

Data (2 population, island model):



PCA results, PC 1:



Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality

Application: testing spatial structures in principal components

PCA results, PC 1:

Moran's *I* test of PC1:



Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Outline

Introduction

Testing spatial structures Moran's Index Mantel's correlation

Correlogram

Multivariate analysis and spatial patterns

Inferring directionality

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Univariate /vs/ multivariate correlation

• Moran's *I* is univariate

- solution: test a few principal components
- problems:
 - does not use all the genetic information
 - which PC to test?
 - correction for multiple testing
- \Rightarrow need for multivariate tests

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Univariate /vs/ multivariate correlation

- Moran's *I* is univariate
- solution: test a few principal components
- problems:
 - does not use all the genetic information
 - which PC to test?
 - correction for multiple testing
- \Rightarrow need for multivariate tests

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Univariate /vs/ multivariate correlation

- Moran's *I* is univariate
- solution: test a few principal components
- problems:
 - · does not use all the genetic information
 - which PC to test?
 - correction for multiple testing
- \Rightarrow need for multivariate tests

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Univariate /vs/ multivariate correlation

- Moran's *I* is univariate
- solution: test a few principal components
- problems:
 - does not use all the genetic information
 - which PC to test?
 - correction for multiple testing
- \Rightarrow need for multivariate tests

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Mantel's correlation: rationale

Correlation between two unfolded distance matrices.



Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Mantel's correlation: definition

Notations:

- $\mathbf{X} = [x_{ij}] \ (\mathbf{X} \in \mathbb{R}^{n \times n})$: genetic distances
- $\mathbf{Y} = [y_{ij}] \ (\mathbf{Y} \in \mathbb{R}^{n \times n})$: geographic distances
- \bar{x} , \bar{y} : means of x and y (excepting diagonals)
- s_x , s_y : standard deviation of x and y (excepting diagonals)

Original definition (unstandardized):

$$z_{\mathsf{M}} = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} x_{ij} y_{ij}$$

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Mantel's correlation: definition

Notations:

- $\mathbf{X} = [x_{ij}] \ (\mathbf{X} \in \mathbb{R}^{n \times n})$: genetic distances
- $\mathbf{Y} = [y_{ij}] \ (\mathbf{Y} \in \mathbb{R}^{n \times n})$: geographic distances
- \bar{x} , \bar{y} : means of x and y (excepting diagonals)
- s_x , s_y : standard deviation of x and y (excepting diagonals)

Standardized coefficient (true correlation):

$$r_{\mathsf{M}} = \frac{1}{d-1} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (\frac{x_{ij} - \bar{x}}{s_x}) (\frac{y_{ij} - \bar{y}}{s_y})$$

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Testing Mantel correlation

- compute $z_{\mbox{\scriptsize M}}$ or $r_{\mbox{\scriptsize M}}$ from the data
- permute randomly the rows and columns of one matrix, recompute the test statistic (i.e., under *H*₀: "no correlation")
- repeat this operation many times to generate a reference distribution
- compare initial value to the reference distribution to get a p-value.

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Testing Mantel correlation

- compute $z_{\mbox{M}}$ or $r_{\mbox{M}}$ from the data
- permute randomly the rows and columns of one matrix, recompute the test statistic (i.e., under *H*₀: "no correlation")
- repeat this operation many times to generate a reference distribution
- compare initial value to the reference distribution to get a p-value.

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Testing Mantel correlation

- compute z_{M} or r_{M} from the data
- permute randomly the rows and columns of one matrix, recompute the test statistic (i.e., under *H*₀: "no correlation")
- repeat this operation many times to generate a reference distribution
- compare initial value to the reference distribution to get a p-value.

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Testing Mantel correlation

- compute z_{M} or r_{M} from the data
- permute randomly the rows and columns of one matrix, recompute the test statistic (i.e., under *H*₀: "no correlation")
- repeat this operation many times to generate a reference distribution
- compare initial value to the reference distribution to get a p-value.

Testing spatial structures ○○○○○○○ ○○○○○● ○○○○○ Multivariate analysis and spatial patterns

Inferring directionality 00000

Application: testing spatial structures

Data (2 population, island model):



Mantel test:



Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Outline

Introduction

Testing spatial structures

Moran's Index Mantel's correlation Correlogram

Multivariate analysis and spatial patterns

Inferring directionality

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Spatial dependence at different scales

Positive autocorrelation can indicate different patterns



 \Rightarrow How to detect the scale of spatial structures?

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Spatial dependence at different scales

Positive autocorrelation can indicate different patterns



 \Rightarrow How to detect the scale of spatial structures?

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Correlogram: rationale

Compute spatial autocorrelation (Moran's I, Mantel's $r_{\mbox{M}})$ for different distance classes.

- define several classes of geographic distances
- for each class, define connectivity between locations
- set spatial weights of non-connected locations to zero
- compute spatial autocorrelation for each class

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Correlogram: rationale

Compute spatial autocorrelation (Moran's ${\it I}$, Mantel's ${\it r}_{\rm M})$ for different distance classes.

- define several classes of geographic distances
- for each class, define connectivity between locations
- set spatial weights of non-connected locations to zero
- compute spatial autocorrelation for each class

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Correlogram: rationale

Compute spatial autocorrelation (Moran's I, Mantel's $r_{\mbox{M}})$ for different distance classes.

- define several classes of geographic distances
- for each class, define connectivity between locations
- set spatial weights of non-connected locations to zero
- compute spatial autocorrelation for each class

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Correlogram: rationale

Compute spatial autocorrelation (Moran's I, Mantel's $r_{\mbox{M}})$ for different distance classes.

- define several classes of geographic distances
- for each class, define connectivity between locations
- set spatial weights of non-connected locations to zero
- compute spatial autocorrelation for each class

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Correlogram: rationale

Compute spatial autocorrelation (Moran's I, Mantel's $r_{\mbox{M}})$ for different distance classes.

- define several classes of geographic distances
- for each class, define connectivity between locations
- set spatial weights of non-connected locations to zero
- compute spatial autocorrelation for each class

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Correlogram: rationale



Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Application: describing spatial structures

Data (2 population, island model):

Mantel correlogram:



Testing spatial structures 000000000 000000 00000 Multivariate analysis and spatial patterns

Inferring directionality 00000

Outline

Introduction

Testing spatial structures

Moran's Index Mantel's correlation Correlogram

Multivariate analysis and spatial patterns

Inferring directionality

Testing spatial structures 000000000 000000 00000 Multivariate analysis and spatial patterns •••••••• Inferring directionality 00000

Mapping principal components

Maps of the three first principal components of PCA.



Is this appropriate?

Testing spatial structures 000000000 000000 000000 Multivariate analysis and spatial patterns •••••••• Inferring directionality 00000

Mapping principal components

Maps of the three first principal components of PCA.



Is this appropriate?

Testing spatial structures 000000000 0000000 000000 Inferring directionality 00000

Rationale of multivariate analyses (recall)

Finding the directions summarizing best the genetic diversity.



Spatial information is not taken into account.

Testing spatial structures 000000000 0000000 000000 Inferring directionality 00000

Rationale of multivariate analyses (recall)

Finding the directions summarizing best the genetic diversity.



Spatial information is not taken into account.

Testing spatial structures 000000000 000000 000000 Multivariate analysis and spatial patterns

Inferring directionality 00000

Using spatial information

- usual multivariate analyses ignore spatial information
- they may reveal obvious spatial structures, but overlook finer patterns

 \Rightarrow need for taking spatial information into account

Testing spatial structures 000000000 000000 000000 Multivariate analysis and spatial patterns

Inferring directionality 00000

Using spatial information

- usual multivariate analyses ignore spatial information
- they may reveal obvious spatial structures, but overlook finer patterns

 \Rightarrow need for taking spatial information into account

Testing spatial structures 000000000 000000 000000 Multivariate analysis and spatial patterns

Inferring directionality 00000

Using spatial information

- usual multivariate analyses ignore spatial information
- they may reveal obvious spatial structures, but overlook finer patterns

 \Rightarrow need for taking spatial information into account

Testing spatial structures 000000000 0000000 000000 Multivariate analysis and spatial patterns

Inferring directionality 00000

Spatial Principal Component Analysis (sPCA): rationale

Principal components \mathbf{Xu} should:

- display variability $\Rightarrow \max var(\mathbf{Xu})$
- display positive autocorrelation $\Rightarrow \max I(\mathbf{Xu})$
- (or) display negative autocorrelation \Rightarrow min $I(\mathbf{Xu})$

Testing spatial structures 000000000 0000000 000000 Multivariate analysis and spatial patterns

Inferring directionality 00000

Spatial Principal Component Analysis (sPCA): rationale

Principal components \mathbf{Xu} should:

- display variability $\Rightarrow \max var(\mathbf{Xu})$
- display positive autocorrelation $\Rightarrow \max I(\mathbf{Xu})$
- (or) display negative autocorrelation \Rightarrow min $I(\mathbf{Xu})$
Testing spatial structures 000000000 0000000 000000 Multivariate analysis and spatial patterns

Inferring directionality 00000

Spatial Principal Component Analysis (sPCA): rationale

Principal components \mathbf{Xu} should:

- display variability $\Rightarrow \max var(\mathbf{Xu})$
- display positive autocorrelation $\Rightarrow \max I(\mathbf{Xu})$
- (or) display negative autocorrelation \Rightarrow min $I(\mathbf{Xu})$

Formalization of the problem

Notations:

- X: matrix of centered (possibly scaled) allele frequencies $(\mathbf{X} \in \mathbb{R}^{n \times p})$
- L: matrix of spatial weights $(\mathbf{L} \in \mathbb{R}^{n imes n})$
- u: principal axis made of p loadings ($\mathbf{u} \in \mathbb{R}^p$, $\|\mathbf{u}\|^2 = 1$)
- Xu: principal component The criterion of the sPCA is:

Additional constraints:

•
$$\|\mathbf{u}\|^2 = 1$$

successive axes must be orthogonal

Formalization of the problem

Notations:

- X: matrix of centered (possibly scaled) allele frequencies $(\mathbf{X} \in \mathbb{R}^{n \times p})$
- L: matrix of spatial weights $(\mathbf{L} \in \mathbb{R}^{n imes n})$
- u: principal axis made of p loadings ($\mathbf{u} \in \mathbb{R}^p$, $\|\mathbf{u}\|^2 = 1$)
- Xu: principal component The criterion of the sPCA is:

$$\begin{array}{cccc} g: & \mathbb{R}^{n \times p} \times \mathbb{R}^{n \times n} \times \mathbb{R}^p & \longrightarrow & \mathbb{R} \\ & & (\mathbf{X}, \mathbf{L}, \mathbf{u}) & \longmapsto & g(\mathbf{X}, \mathbf{L}, \mathbf{u}) = \mathsf{var}(\mathbf{X}\mathbf{u})I(\mathbf{X}\mathbf{u}) \end{array}$$

Additional constraints:

•
$$\|\mathbf{u}\|^2 = 1$$

successive axes must be orthogonal

Testing spatial structures 000000000 000000 00000 Multivariate analysis and spatial patterns

Inferring directionality 00000

Solution

• It can be shown that:

$$g(\mathbf{X}, \mathbf{L}, \mathbf{u}) = \frac{1}{2n} \mathbf{u}^T \mathbf{X}^T (\mathbf{L} + \mathbf{L}^T) \mathbf{X} \mathbf{u}$$

where $\mathbf{X}^{\mathit{T}}(\mathbf{L}+\mathbf{L}^{\mathit{T}})\mathbf{X}$ is symmetric.

• The extremums of $g(\mathbf{X}, \mathbf{L}, \mathbf{u})$ are then given by the eigenanalysis of:

$$\frac{1}{2n}\mathbf{X}^T(\mathbf{L}+\mathbf{L}^T)\mathbf{X}$$

• We obtain a set of orthonormal eigenvectors $\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$ associated to r non-null eigenvalues $\{\lambda_1, \cdots, \lambda_r\}$ (with $\lambda_i > \lambda_{i+1}$, i = 1, ..., r - 1) verifying:

$$\lambda_r = \mathsf{var}(\mathbf{X}\mathbf{u}_r) I(\mathbf{X}\mathbf{u}_r) \leq g(\mathbf{X},\mathbf{L},\mathbf{u}_i) \leq \mathsf{var}(\mathbf{X}\mathbf{u}_1) I(\mathbf{X}\mathbf{u}_1) = \lambda_1$$

Testing spatial structures 000000000 000000 00000 Multivariate analysis and spatial patterns

Inferring directionality 00000

Solution

• It can be shown that:

$$g(\mathbf{X}, \mathbf{L}, \mathbf{u}) = \frac{1}{2n} \mathbf{u}^T \mathbf{X}^T (\mathbf{L} + \mathbf{L}^T) \mathbf{X} \mathbf{u}$$

where $\mathbf{X}^{\mathit{T}}(\mathbf{L}+\mathbf{L}^{\mathit{T}})\mathbf{X}$ is symmetric.

• The extremums of $g(\mathbf{X}, \mathbf{L}, \mathbf{u})$ are then given by the eigenanalysis of:

$$\frac{1}{2n}\mathbf{X}^T(\mathbf{L}+\mathbf{L}^T)\mathbf{X}$$

We obtain a set of orthonormal eigenvectors {u₁, · · · , u_r} associated to r non-null eigenvalues {λ₁, · · · , λ_r} (with λ_i > λ_{i+1}, i = 1, ..., r - 1) verifying:

$$\lambda_r = \mathsf{var}(\mathbf{X}\mathbf{u}_r) I(\mathbf{X}\mathbf{u}_r) \leq g(\mathbf{X}, \mathbf{L}, \mathbf{u}_i) \leq \mathsf{var}(\mathbf{X}\mathbf{u}_1) I(\mathbf{X}\mathbf{u}_1) = \lambda_1$$

Testing spatial structures 000000000 000000 00000 Multivariate analysis and spatial patterns

Inferring directionality 00000

Solution

• It can be shown that:

$$g(\mathbf{X}, \mathbf{L}, \mathbf{u}) = \frac{1}{2n} \mathbf{u}^T \mathbf{X}^T (\mathbf{L} + \mathbf{L}^T) \mathbf{X} \mathbf{u}$$

where $\mathbf{X}^{\mathit{T}}(\mathbf{L}+\mathbf{L}^{\mathit{T}})\mathbf{X}$ is symmetric.

• The extremums of $g(\mathbf{X}, \mathbf{L}, \mathbf{u})$ are then given by the eigenanalysis of:

$$\frac{1}{2n}\mathbf{X}^{T}(\mathbf{L}+\mathbf{L}^{T})\mathbf{X}$$

We obtain a set of orthonormal eigenvectors {u₁, · · · , u_r} associated to r non-null eigenvalues {λ₁, · · · , λ_r} (with λ_i > λ_{i+1}, i = 1, ..., r − 1) verifying:

$$\lambda_r = \mathsf{var}(\mathbf{X}\mathbf{u}_r) I(\mathbf{X}\mathbf{u}_r) \leq g(\mathbf{X},\mathbf{L},\mathbf{u}_i) \leq \mathsf{var}(\mathbf{X}\mathbf{u}_1) I(\mathbf{X}\mathbf{u}_1) = \lambda_1$$

Testing spatial structures

Multivariate analysis and spatial patterns

Inferring directionality 00000

Spatial Principal Component Analysis (sPCA): outputs



Testing spatial structures 000000000 000000 00000 Multivariate analysis and spatial patterns 0000000000

Inferring directionality 00000

Spatial Principal Component Analysis (sPCA): example

Chamois (Rupicapra rupicapra) in the Bauges mountains.



What is the genetic structure of this population (reproduction units)?

Testing spatial structures 000000000 000000 00000 Multivariate analysis and spatial patterns 0000000000

Inferring directionality 00000

Spatial Principal Component Analysis (sPCA): example

Chamois (Rupicapra rupicapra) in the Bauges mountains.



What is the genetic structure of this population (reproduction units)?

Testing spatial structures 000000000 0000000 000000 Multivariate analysis and spatial patterns

Inferring directionality 00000

Spatial Principal Component Analysis (sPCA): example

PCA results: no spatial structure, 4 outliers



Testing spatial structures 000000000 0000000 000000 Multivariate analysis and spatial patterns 000000000

Inferring directionality 00000

Spatial Principal Component Analysis (sPCA): example

sPCA results: several spatial structures (confirmed by field observations).



Testing spatial structures 000000000 000000 00000 Multivariate analysis and spatial patterns

Inferring directionality

Outline

Introduction

Testing spatial structures

Moran's Index Mantel's correlation Correlogram

Multivariate analysis and spatial patterns

Inferring directionality

Testing spatial structures 000000000 0000000 000000 Multivariate analysis and spatial patterns 000000000

Inferring directionality •0000

Multivariate method do not infer directionality

Principal Components have no direction.



(Cavalli-Sforza et al. 1993, *Science*) How to infer migrations?

Testing spatial structures 000000000 0000000 000000 Multivariate analysis and spatial patterns

Inferring directionality •0000

Multivariate method do not infer directionality

Principal Components have no direction.



(Cavalli-Sforza et al. 1993, *Science*) How to infer migrations?

Testing spatial structures 00000000 000000 00000 Multivariate analysis and spatial patterns 000000000

Inferring directionality 0000

The 'bottleneck' effect



Multivariate analysis and spatial patterns

Inferring directionality

Effects of successive bottlenecks

Progressive loss of genetic diversity within populations.



Testing spatial structures 000000000 000000 000000 Multivariate analysis and spatial patterns

Inferring directionality

Expected patterns of migrations

Diversity within population decreases with distance from the origin



Multivariate analysis and spatial patterns 000000000

Inferring directionality

Patterns of migrations: example

Reconstructing the spread of Malaria worlwide.



(Tanabe et al. 2010, Current Biology)

Plasmodium falciparum accompanied Human migrations.